

University of Groningen

## Inhibition-augmented trainable COSFIRE filters for keypoint detection and object recognition

Guo, Jiapan; Shi, Chenyu; Azzopardi, George; Petkov, Nicolai

*Published in:*  
Machine Vision and Applications

*DOI:*  
[10.1007/s00138-016-0777-3](https://doi.org/10.1007/s00138-016-0777-3)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2016

[Link to publication in University of Groningen/UMCG research database](#)

### *Citation for published version (APA):*

Guo, J., Shi, C., Azzopardi, G., & Petkov, N. (2016). Inhibition-augmented trainable COSFIRE filters for keypoint detection and object recognition. *Machine Vision and Applications*, 1-15.  
<https://doi.org/10.1007/s00138-016-0777-3>

### **Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### **Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Inhibition-augmented trainable COSFIRE filters for keypoint detection and object recognition

Jiapan Guo<sup>1</sup> · Chenyu Shi<sup>1</sup> · George Azzopardi<sup>2</sup> · Nicolai Petkov<sup>1</sup>

Received: 20 January 2016 / Revised: 28 April 2016 / Accepted: 5 May 2016 / Published online: 27 May 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** The shape and meaning of an object can radically change with the addition of one or more contour parts. For instance, a T-junction can become a crossover. We extend the COSFIRE trainable filter approach which uses a positive prototype pattern for configuration by adding a set of negative prototype patterns. The configured filter responds to patterns that are similar to the positive prototype but not to any of the negative prototypes. The configuration of such a filter comprises selecting given channels of a bank of Gabor filters that provide excitatory or inhibitory input and determining certain blur and shift parameters. We compute the response of such a filter as the excitatory input minus a fraction of the maximum of inhibitory inputs. We use three applications to demonstrate the effectiveness of inhibition: the exclusive detection of vascular bifurcations (i.e., without crossovers) in retinal fundus images (DRIVE data set), the recognition of architectural and electrical symbols (GREC'11 data set) and the recognition of handwritten digits (MNIST data set).

**Keywords** Inhibition · Trainable COSFIRE filter · Keypoints detection · Object recognition

✉ Chenyu Shi  
c.shi@rug.nl

Jiapan Guo  
j.guo@rug.nl

George Azzopardi  
george.azzopardi@um.edu.mt

Nicolai Petkov  
n.petkov@rug.nl

<sup>1</sup> Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Nijenborgh 9, 9747 AG Groningen, Netherlands

<sup>2</sup> Intelligent Computer Systems, University of Malta, Msida MSD 2080, Malta

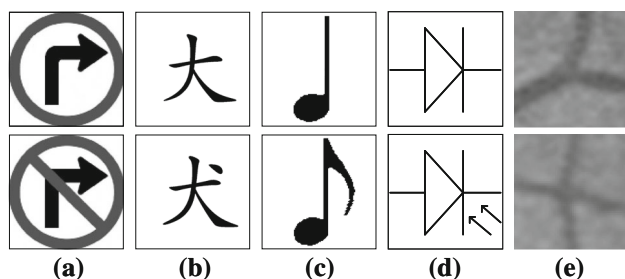
## 1 Introduction

Recently, a novel trainable filter for object recognition has been proposed in [5]. It is called combination of shifted filter responses or COSFIRE for brevity. A COSFIRE filter is configured to be selective for a given local pattern by extracting from that pattern characteristic properties of contour parts (such as orientation) and their geometrical arrangement. COSFIRE filters were demonstrated to be effective for detection of local patterns (keypoints) and recognition of objects and achieve very good performance in various applications [4, 6, 8, 19, 47, 49, 50]. They were also used in a multilayer hierarchical approach [6].

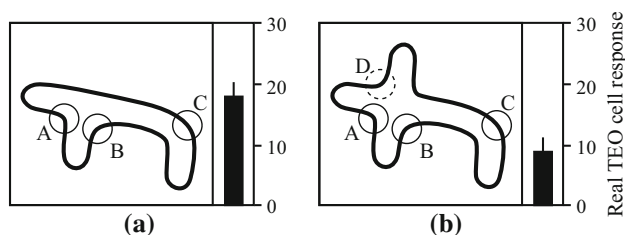
Figure 1 shows some examples where a COSFIRE filter of the type proposed in [5] may, however, not perform very well. COSFIRE filters that are configured to be selective for the patterns shown in the images in the top row of Fig. 1 also give strong responses to the images in the bottom row of Fig. 1. This is because all contour parts of a pattern in the top row are present, in the preferred arrangements, in the corresponding image shown in the bottom row of Fig. 1. The presence of additional contour parts, such as the diagonal bar in Fig. 1a (bottom) or the extra stroke in Fig. 1b (bottom), does not have influence on the response of the filter.

The COSFIRE method [5] was inspired by a specific type of shape-selective neuron in area V4 of visual cortex. This method, however, relies on contour parts that provide only excitatory inputs. This means that every involved contour part detector contributes to enhance the response of a COSFIRE filter.

There is neurophysiological evidence, however, that neurons in different layers of the visual cortex receive also inhibitory inputs [21]. For instance, neurons in the lateral geniculate nucleus (LGN) have center-surround receptive fields which have been modeled by difference-of-Gaussians



**Fig. 1** Examples of pairs of patterns that COSFIRE filters of the type proposed in [5] are not able to distinguish. **a** Two traffic signs which have the opposite messages: permission and prohibition of turning right. **b** Two Chinese characters that are translated into English as “big” and “dog.” **c** Two music notes: quarter and eighth. **d** Two electrical symbols: normal and light-emitting diodes. **e** Bifurcations and crossovers in retinal fundus images. A COSFIRE filter that is trained to detect the upper pattern in a pair of two will give a strong response to the lower pattern too



**Fig. 2** Selectivity of a shape-selective neuron in the posterior inferotemporal cortex [13]. **a** The curvatures marked with *circles* evoke excitation of the concerned cell, while **b** the curvature marked with a *dashed circle* inhibits the activation of the cell. The bars specify the strength of the response

(DoG) operators. A center-on DoG has an excitatory central region with an inhibitory surround. Similarly, simple cells in area V1, whose properties provided the inspiration for Gabor filters [16, 23], derivative of Gaussians [18] and CORF [3, 7] filters, have receptive fields that consist of inhibitory and excitatory regions. Non-classical receptive field inhibition in orientation-selective visual neurons provided the inspiration for surround inhibition in orientation-selective filters [42]. It has been shown to improve contour detection by suppressing responses to textured regions. Moreover, shape-selective neurons of the type studied in [13], located in the posterior inferotemporal cortex, respond to complex shapes that are formed by a number of convex and concave curvatures with a certain geometrical arrangement. The presence of some specific curvature elements can inhibit the response of such a neuron. Figure 2 shows the response of a TEO neuron, studied in [13], which is excited by the encircled curvatures A, B and C, but is inhibited by the dashed encircled curvature D. The bar plots indicate the responses to the stimuli. Inhibition is also thought to increase the selectivity of neurons [46].

Inhibition is an important phenomenon in the brain. It facilitates sparseness in the representation of information

that may result in an increase in the storage capacity and a higher number of patterns that can be discriminated [45]. End-stopped cells [12, 22] in area V1 of visual cortex are another example.

In this work, we add inhibition to COSFIRE filters in order to increase their discrimination ability. The inhibition that we propose is learned in an automatic configuration process. We configure an inhibition-augmented COSFIRE filter by using two different types of prototype patterns, namely one positive pattern and one or more negative pattern(s), in order to extract excitatory and inhibitory contour parts, respectively. Such a filter can effectively detect patterns that are equivalent or similar to the positive prototype, but does not respond to the negative prototype(s).

The proposed inhibition-augmented filters can be used in keypoint detection and object recognition. A large body of work has been done in these areas, and many methods have been proposed [9, 10, 15, 20, 24, 28, 34–39, 56, 59]. The Hessian detector [10] and the Harris detector [20], for instance, detect points of interest and are invariant to rotation but not so much to scale. Scaling invariances of these two operators can be achieved by applying them in a Laplacian of Gaussian scale space [34], resulting in the so-called Hessian–Laplace and Harris–Laplace detectors [36]. A point of interest can be described by some local keypoint descriptors, such as the scale-invariant feature transform (SIFT) [35], the histogram of oriented gradients (HOG) [15], the image descriptor GIST [39] and the gradient location and orientation histogram (GLOH) [37]. Other keypoint descriptors include the speeded up robust features (SURF) [9], which is akin to SIFT but faster as it makes efficient use of integral images [56], the texture-based local binary patterns (LBP) [38], textons [24, 59] and the biologically inspired local descriptor (BILD) [60], as well as the rotation invariant feature transform (RIFT) descriptor [28]. None of these methods employs inhibition.

Multiple keypoints can be used to represent bigger and more complex patterns, such as complete objects or scenes. In [32], a bag-of-visual-words approach was proposed to describe an image or a region of interest with a histogram of prototypical keypoints. This method is improved by using spatial pyramids [29] or a random sample consensus algorithm [25]. Other object recognition approaches use hierarchical representations of objects, which have been inspired by the visual processing in the brain. These include the HMAX model [44], the object representation by parts proposed in [17], neural networks [26] and the deep learning approach [30].

These methods require many training examples to configure models of objects of interest. When such detectors and descriptors are trained, only positive examples are considered without the inclusion of inhibition mechanisms. The resulting detectors and descriptors can detect objects that are similar

to positive examples, but may also give strong responses to objects that contain additional contour parts. For instance, the detectors, which are trained by examples shown in the top row of Fig. 1, will give strong responses to objects that are equivalent or similar to the ones shown in the top row of Fig. 1. They will, however, also give strong responses to objects that are equivalent or similar to the ones in the bottom row of Fig. 1. Therefore, it is difficult for these methods to discriminate the pairs of patterns shown in Fig. 1a–f.

The rest of the paper is organized as follows. In Sect. 2, we explain how an inhibition-augmented filter is configured by given positive and negative prototype patterns. In Sect. 3, we demonstrate the effectiveness of the proposed approach in three applications. In Sect. 4, we discuss some aspects of the proposed method, and finally, we draw conclusions in Sect. 5.

## 2 Method

### 2.1 Overview

Figure 3a shows an input image containing a rectangle with a vertical line inside it. Let us consider the two local patterns encircled by a solid and a dashed line, which are shown enlarged in Fig. 3b, c, respectively. The two solid ellipses in Fig. 3b, c surround a line segment that is present in both patterns, while the dashed ellipse surrounds a line segment that is only present in Fig. 3c. We use these two patterns to configure an inhibition-augmented filter that will respond to the pattern shown in Fig. 3b, a line ending, but not to the pattern shown in Fig. 3c, a continuous line.

We consider the line ending and the continuous line shown in Fig. 3b, c as a positive and a negative prototype, respectively. A positive prototype is a local pattern to which the

inhibition-augmented filter to be configured should respond, while a negative prototype is a local pattern to which it should not respond.

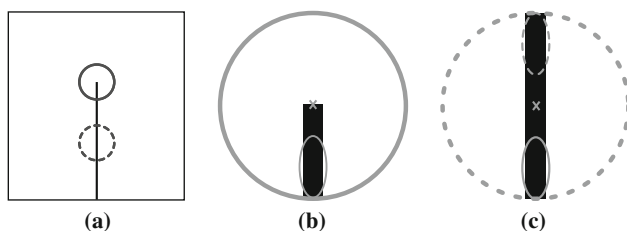
We use the positive and the negative prototypes to configure two COSFIRE filters with the method proposed in [5]. Next, we look for and identify pairs of contour parts with identical properties in the two filters. In Fig. 3, we use a solid ellipse to indicate that the corresponding contour part is an excitatory feature. We use a dashed ellipse to indicate the contour part that is only present in the negative prototype, and therefore, we consider it as an inhibitory feature.

The response of the inhibition-augmented filter is the difference between the *excitatory input* and a fraction of the maximum of the *inhibitory inputs*. The resulting filter will only respond to the patterns that are identical with or similar to the positive prototype, but will not respond to images similar to any of the negative prototypes. This design decision is inspired by the function of a type of shape-selective neuron in posterior inferotemporal cortex.

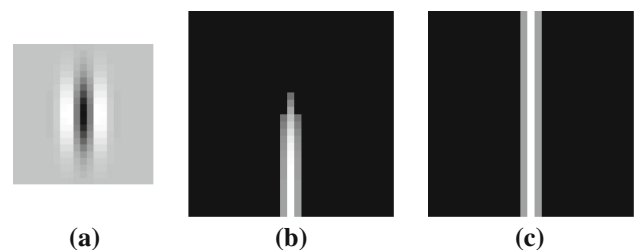
In the next subsections, we elaborate further on the configuration steps mentioned above.

### 2.2 Gabor filters

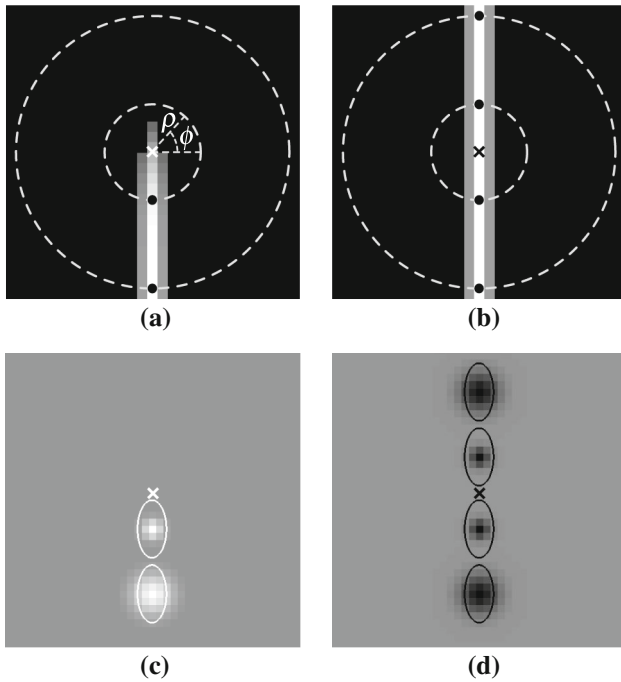
The proposed inhibition-augmented filter uses as input the responses of Gabor filters. We denote by  $g_{\lambda,\theta}(x, y)$  the response of a Gabor filter, which has a preferred wavelength  $\lambda$  and orientation  $\theta$ , to a given input image at location  $(x, y)$ . We threshold the responses of Gabor filters at a given fraction  $t_1$  ( $0 \leq t_1 \leq 1$ ) of the maximum response across all combinations of values  $(\lambda, \theta)$  and all positions  $(x, y)$  in the image. We denote these thresholded response images by  $|g_{\lambda,\theta}(x, y)|_{t_1}$ . Figure 4a shows the intensity map of a Gabor function with a wavelength  $\lambda = 6$  and an orientation  $\theta = 0$ . Figure 4b, c shows the corresponding thresholded response images of this Gabor filter  $|g_{6,0}(x, y)|_{t_1=0.2}$  to the input images in Fig. 3b, c, respectively. Such a filter has other parameters, including spatial aspect ratio, bandwidth and phase offset on which we do not elaborate further here. We refer the interested reader



**Fig. 3** **a** Synthetic input image (of size  $300 \times 300$  pixels). The *solid circle* indicates a positive prototype of interest (a line ending) while the *dashed circle* indicates a negative prototype of interest (a continuous line segment). The images in **(b, c)** show enlargements of the selected positive and negative prototype patterns, respectively. The *gray crosses* in **(b, c)** indicate the center positions of interest, and the ellipses illustrate the orientation and location of the contour parts in the neighborhoods. The *solid ellipses* represent line segments that are present in both prototypes, while the *dashed ellipse* represents a line segment which is only present in the negative prototype



**Fig. 4** **a** Intensity map (of size  $21 \times 21$  pixels) of a symmetric Gabor function with wavelength  $\lambda = 6$  and orientation  $\theta = 0$ . Light and dark regions correspond to positive and negative values of the Gabor function, respectively. **b–c** The thresholded (at  $t_1 = 0.2$ ) Gabor response images (of size  $30 \times 30$  pixels) to Fig. 3b, c, respectively



**Fig. 5** Configuration of two COSFIRE filters as proposed in [5]. Thresholded responses of a symmetric Gabor filter with  $\lambda = 6$  and  $\theta = 0$  to **a** the positive and **b** the negative prototypes are shown in Fig. 3b, c, respectively. The *cross markers* indicate the centers of the filter. We consider the *dashed circles* of given radii (here  $\rho \in \{5, 14\}$ ) around the center of the pattern of interest. The *black dots* indicate the positions of the local maxima of the Gabor responses along these circles. In each such point, we select the Gabor filter which gives this local maximum response. **c, d** Structure of the resulting two COSFIRE filters. The ellipses illustrate the wavelengths and orientations of the selected Gabor filters, and their positions indicate the locations at which the responses of these Gabor filters are taken with respect to the center. The *blobs* represent the blurring functions that are used to provide some spatial tolerance to these positions

to [5, 27, 41] for technical details and to an online implementation.<sup>1</sup>

### 2.3 Configuration of an inhibition-augmented filter

The configuration of an inhibition-augmented filter involves two steps.

In the first step, we configure two separate COSFIRE filters with the method proposed in [5] to be selective for the specified positive and negative prototypes that are shown in Fig. 3b, c, respectively. Figure 5a, b shows the corresponding superimposed thresholded responses of a bank of Gabor filters ( $\theta \in \{0, \pi/8, \dots, 7\pi/8\}$  and  $\lambda \in \{4, 4\sqrt{2}, 6, 6\sqrt{2}\}$ ) to the positive and negative prototypes. In this example, for the configuration of a COSFIRE filter with a given prototype, we consider the Gabor responses along two concentric circles with radii  $\rho \in \{5, 14\}$  pixels around the specified

point of interest. In Fig. 5c, d we illustrate the structures of the resulting selected filters. The size and orientation of an ellipse represent the preferred wavelength  $\lambda$  and orientation  $\theta$  of a Gabor filter that provides input to the COSFIRE filter. The position of its center indicates the location at which we take the concerned Gabor filter response.

We specify a COSFIRE filter by a set of four tuples in which each four tuple represents a Gabor filter and the positions at which its response has to be taken. We denote by  $P_f$  and  $N_f$  the two COSFIRE filters, configured with the patterns shown in Fig. 3b, c, respectively:

$$P_f = \left\{ (\lambda_1 = 6, \theta_1 = 0, \rho_1 = 5, \phi_1 = 3\pi/2), \right. \\ \left. (\lambda_2 = 6, \theta_2 = 0, \rho_2 = 14, \phi_2 = 3\pi/2) \right\}$$

and

$$N_f = \left\{ (\lambda_1 = 6, \theta_1 = 0, \rho_1 = 5, \phi_1 = \pi/2), \right. \\ \left. (\lambda_2 = 6, \theta_2 = 0, \rho_2 = 5, \phi_2 = 3\pi/2), \right. \\ \left. (\lambda_3 = 6, \theta_3 = 0, \rho_3 = 14, \phi_3 = \pi/2), \right. \\ \left. (\lambda_4 = 6, \theta_4 = 0, \rho_4 = 14, \phi_4 = 3\pi/2) \right\}$$

In the second step, we form a new set  $S_f$  by selecting tuples from the sets  $P_f$  and  $N_f$  as follows. We include all tuples from the set  $P_f$  in the new set  $S_f$  and add a new parameter  $\delta = +1$  to indicate that the corresponding Gabor responses of such tuples provide excitatory input to the inhibition-augmented filter. We define a dissimilarity function, which we denote by  $d(P_f^i, N_f^j)$ , of the distance between the locations indicated by the  $i$ th tuple in the set  $P_f$  and the  $j$ th tuple in the set  $N_f$ :

$$d(P_f^i, N_f^j) = \begin{cases} 1, & \text{if } D > \zeta \\ 0, & \text{otherwise} \end{cases}$$

$$D = \sqrt{(\rho_i \cos \phi_i - \rho_j \cos \phi_j)^2 + (\rho_i \sin \phi_i - \rho_j \sin \phi_j)^2} \quad (1)$$

where  $D$  is the Euclidean distance between the polar coordinates  $(\rho_i, \phi_i)$  of tuple  $i$  in the positive set  $P_f$  and the polar coordinates  $(\rho_j, \phi_j)$  of tuple  $j$  in the negative set  $N_f$ .  $\zeta$  is the threshold, and we provide further details on the selection of its value in Sect. 2.5.

We compute the pairwise dissimilarity values between one tuple  $N_f^j$  from  $N_f$  and all tuples from  $P_f$ . If  $N_f^j$  is dissimilar to all tuples in  $P_f$ , we include it to the new set  $S_f$  and add a tag  $\delta = -1$ , which indicates that the corresponding Gabor response provides an inhibitory input. We repeat the above procedure for each tuple in set  $N_f$ . With this process, we ensure that a line segment that is present in both the positive and the negative prototypes in roughly the same position gives an excitatory input. On the other hand, a line segment that is only present in the negative prototype, i.e., it does

<sup>1</sup> <http://matlabserver.cs.rug.nl>.



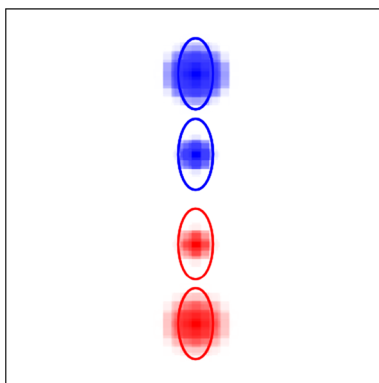
not overlap with a line segment in the positive prototype, provides an inhibitory input.

For the above example, we include the two tuples in set  $P_f$ , which are illustrated by the two ellipses in Fig. 5c, in the new set  $S_f$ . We add to each of these two tuples a tag  $\delta = +1$  to indicate that they provide excitatory input to the inhibition-augmented filter. These two tuples are also present in set  $N_f$ . Then, we include in  $S_f$  the other two tuples from  $N_f$  indicated by the two ellipses at the top of Fig. 5d with a tag  $\delta = -1$  as we do not find any matches in  $P_f$ . For the above example, this method results in the following set  $S_f$ :

$$S_f = \left\{ \begin{array}{l} (\lambda_1 = 6, \theta_1 = 0, \rho_1 = 5, \phi_1 = 3\pi/2, \delta_1 = +1), \\ (\lambda_2 = 6, \theta_2 = 0, \rho_2 = 14, \phi_2 = 3\pi/2, \delta_2 = +1), \\ (\lambda_3 = 6, \theta_3 = 0, \rho_3 = 5, \phi_3 = \pi/2, \delta_3 = -1), \\ (\lambda_4 = 6, \theta_4 = 0, \rho_4 = 14, \phi_4 = \pi/2, \delta_4 = -1) \end{array} \right\}$$

Figure 6 shows the structure of the resulting inhibition-augmented filter that is represented by the set  $S_f$ . The red ellipses indicate Gabor filters that provide excitatory input, and the blue ellipses indicate Gabor filters that provide inhibitory input to the inhibition-augmented filter at hand.

For example, the second tuple in  $S_f$  ( $\lambda_2 = 6, \theta_2 = 0, \rho_2 = 14, \phi_2 = 3\pi/2, \delta_2 = +1$ ) corresponds to the ellipse in the bottommost of Fig. 6. It describes a line segment with a width of ( $\lambda_2/2 =$ ) 3 pixels in a vertical ( $\theta_2 = 0$ ) orientation at a position of ( $\rho_2 =$ ) 14 pixels to the bottom ( $\phi_2 = 3\pi/2$ ) of the point of interest. This tuple provides excitatory ( $\delta_2 = +1$ ) input to the inhibition-augmented filter. On the other hand, the last tuple in  $S_f$  ( $\lambda_4 = 6, \theta_4 = 0, \rho_4 = 14, \phi_4 = \pi/2, \delta_4 = -1$ ) corresponds to the topmost ellipse in Fig. 6. It describes a similar line segment at a position of ( $\rho_2 =$ ) 14 pixels to the top ( $\phi_4 = \pi/2$ ) of the point of interest and provides inhibitory ( $\delta_4 = -1$ ) input to the filter.



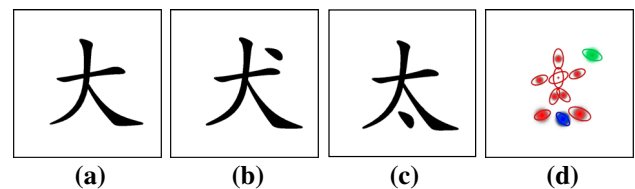
**Fig. 6** Structure of an inhibition-augmented filter. The four ellipses indicate the responses of four Gabor filters with the parameter values specified by the set  $S_f$ . The two red ellipses at the bottom represent the excitatory input to this inhibition-augmented filter, while the two blue ellipses at the top represent the inhibitory input (color figure online)

## 2.4 Configuration with multiple negative prototypes

In the above example, we configured an inhibition-augmented filter to be selective for line endings by using one positive and one negative prototype pattern. In practice, however, a positive pattern may be contained within multiple other patterns, and thus, we may need multiple negative examples.

Figure 7a–c shows an example of three similar Chinese letters that have completely different meanings and are translated into English as “big,” “dog” and “extremely,” respectively. The letter in Fig. 7a is also present in Fig. 7b, c, but accompanied with additional strokes. Next, we demonstrate how we configure an inhibition-augmented filter with more than one negative prototype pattern. Here, we use the letter image in Fig. 7a as our positive pattern of interest from which we extract contour parts that provide excitatory input to the resulting filter. The letter images in Fig. 7b, c are used as negative prototype patterns from which we determine inhibitory contour parts.

First, we configure a filter  $P_f$  for the positive prototype pattern in Fig. 7a as proposed in [5] that results in only excitatory inputs. For this example, we consider three values of the radius  $\rho$  ( $\rho = \{0, 15, 33\}$ ) and we apply a bank of Gabor filters with four wavelengths ( $\lambda \in \{8, 8\sqrt{2}, 16\}$ ) and eight orientations ( $\theta \in \{\frac{\pi i}{8} \mid i = 0 \dots 7\}$ ). Then, we use the procedure proposed in [5] to apply the filter  $P_f$  to both the negative prototype patterns in Fig. 7b, c. For each negative pattern, we determine the location at which the maximum response is achieved by the filter  $P_f$ . We take the patterns from Fig. 7b, c that surround these locations and use them to configure two COSFIRE filters, which we denote by  $N_{f_1}$  and  $N_{f_2}$ , respectively. Finally, we form a new set  $S_{\text{big}}$  by selecting appropriate tuples from  $P_f$ ,  $N_{f_1}$  and  $N_{f_2}$  as follows. We include all tuples from set  $P_f$  in the new set  $S_{\text{big}}$  with a tag  $\delta = +1$  and compute the dissimilarity values between the locations of the tuples in  $N_{f_i}$  (here  $i = 1, 2$ ) and those in set  $P_f$  by the method described in Sect. 2.3. The tuples in  $N_{f_1}$  and  $N_{f_2}$  that are not similar to any of the tuples in  $P_f$  are added to  $S_{\text{big}}$  and marked as inhibitory parts with tags



**Fig. 7** a–c Images of three similar Chinese letters (of size  $120 \times 120$  pixels). d Structure of the resulting filter  $S_{\text{big}}$  that is configured by using a the positive prototype and b, c the two negative prototypes. The red ellipses represent the preferred wavelengths and orientations of the Gabor filter responses that provide excitatory input to the concerned filter  $S_{\text{big}}$ , while the blue and green ellipses represent the Gabor filter responses provide inhibitory input (color figure online)

$\delta = -1$  and  $\delta = -2$ , respectively. These two different negative tags indicate that inhibitory contour parts are extracted from two separate negative patterns.

$$S_{\text{big}} = \left\{ \begin{array}{l} (\lambda_1 = 10, \theta_1 = \frac{\pi}{4}, \rho_1 = 15, \phi_1 = \frac{25\pi}{16}, \delta_1 = +1), \\ \dots \\ (\lambda_9 = 14, \theta_9 = \frac{3\pi}{8}, \rho_9 = 33, \phi_9 = \frac{27\pi}{16}, \delta_9 = +1), \\ (\lambda_{10} = 12, \theta_{10} = \frac{3\pi}{8}, \rho_{10} = 33, \phi_{10} = \frac{3\pi}{16}, \delta_{10} = -1), \\ (\lambda_{11} = 10, \theta_{11} = \frac{\pi}{4}, \rho_{11} = 33, \phi_{11} = \frac{3\pi}{2}, \delta_{11} = -2) \end{array} \right\}$$

Figure 7d shows the resulting structure of the inhibition-augmented filter  $S_{\text{big}}$ , in which the red ellipses indicate the tuples of the filter that provide excitatory input to the inhibition-augmented filter, while the blue and green ellipses indicate the tuples that provide inhibitory input.

## 2.5 Application of an inhibition-augmented COSFIRE filter

In the following, we first explain how we blur and shift the responses of the involved Gabor filters, and then, we describe the functions that we use to compute the collective excitatory input, the various collections of inhibitory inputs and the ultimate filter output.

### 2.5.1 Blurring and shifting Gabor filter responses

We blur the Gabor filter responses in order to allow for some tolerance in the positions at which their responses are taken. We define the blurring operation as the weighted maximum of local Gabor filter responses. For weighting, we use a Gaussian function  $G_\sigma(x, y)$ , the standard deviation  $\sigma$  of which is a linear function of the distance  $\rho$  from the center of the COSFIRE filter:

$$\sigma = \sigma_0 + \alpha\rho \quad (2)$$

where  $\sigma_0$  and  $\alpha$  are constants. The choice of the linear function in Eq. 2 is advocated for more detail in [5]. For  $\alpha > 0$ , the tolerance to the positions of the considered contour parts increases with an increasing distance  $\rho$  from the center of the concerned COSFIRE filter. We use values of  $\alpha$  between 0 and 2, depending on the application.

Then, we shift all blurred Gabor filter responses so that they meet at the support center of the inhibition-augmented filter. This is achieved by shifting the blurred responses of a Gabor filter  $(\lambda_i, \theta_i)$  by a distance  $\rho_i$  in the direction opposite to  $\phi_i$ . In polar coordinates, the shift vector is specified by  $(\rho_i, \phi_i + \pi)$ . In Cartesian coordinates, it is  $(\Delta x_i, \Delta y_i)$  where  $\Delta x_i = -\rho_i \cos \phi_i$ , and  $\Delta y_i = -\rho_i \sin \phi_i$ . We denote by  $s_{\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i}(x, y)$  the blurred and shifted thresholded response of a Gabor filter in position  $(x, y)$  that is specified by the  $i$ th tuple  $(\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i)$  in the set  $S_f$ :

$$s_{\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i}(x, y) \stackrel{\text{def}}{=} \max_{x', y'} \{ |g_{\lambda_i, \theta_i}(x - x' - \Delta x_i, y - y' - \Delta y_i)|_{t_1} G_\sigma(x', y') \} \quad (3)$$

where  $-3\sigma \leq x', y' \leq 3\sigma$ .

In order to prevent interference of inhibitory and excitatory parts of the filter, we restrict  $\zeta$  (in Eq. 1) to be three times the maximum standard deviation of any pair of tuples in  $P_f$  and  $N_f$ .

### 2.5.2 Response of an inhibition-augmented COSFIRE filter

We denote by  $r_{S_f}(x, y)$  the response of an inhibition-augmented COSFIRE filter which we define as the difference between excitatory response  $r_{S_f^+}(x, y)$  and a fraction of the maximum of the inhibitory responses  $r_{S_f^-j}(x, y)$ .

$$r_{S_f}(x, y) \stackrel{\text{def}}{=} |r_{S_f^+}(x, y) - \eta \max_{j=1}^n \{r_{S_f^-j}(x, y)\}|_{t_3} \quad (4)$$

where  $S_f^+ = \{(\lambda_i, \theta_i, \rho_i, \phi_i) \mid \forall (\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i) \in S_f, \delta_i = +1\}$ ,  $S_f^{-j} = \{(\lambda_i, \theta_i, \rho_i, \phi_i) \mid \forall (\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i) \in S_f, \delta_i = -j\}$ ,  $n = \max |\delta_i|$ ,  $\eta$  is a coefficient that we call inhibition factor and  $|\cdot|_{t_3}$  stands for thresholding the response at a fraction  $t_3$  of its maximum across all image coordinates  $(x, y)$ .

We denote by  $r_{S_f^+}$  and  $r_{S_f^-}$ , the weighted geometric means of all the blurred and shifted responses of the Gabor filters  $s_{\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i}(x, y)$  that correspond to the contour parts described by  $S_f^+$  and  $S_f^-$ :

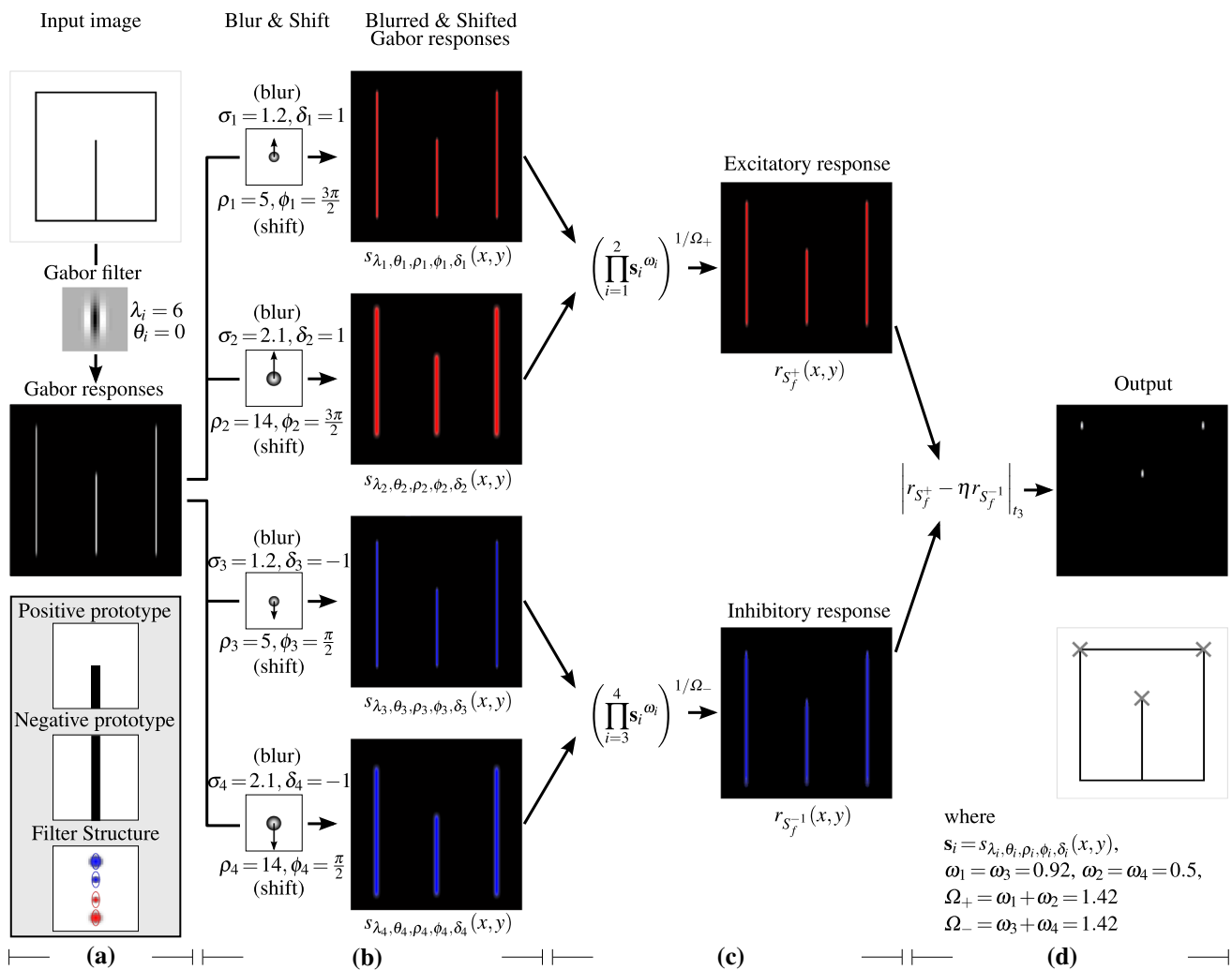
$$r_{S_f^\delta}(x, y) \stackrel{\text{def}}{=} \left| \left( \prod_{i=1}^{|S_f^\delta|} (s_{\lambda_i, \theta_i, \rho_i, \phi_i, \delta_i}(x, y))^{\omega_i} \right)^{1/\sum_{i=1}^{|S_f^\delta|} \omega_i} \right|_{t_2} \quad (5)$$

$$\omega_i = \exp \left( -\frac{\rho_i^2}{2\sigma^2} \right), \quad 0 \leq t_2 \leq 1$$

where  $|\cdot|_{t_2}$  stands for thresholding the response at a fraction  $t_2$  of its maximum across all image coordinates  $(x, y)$ . For  $1/\sigma' = 0$ , the computation of the COSFIRE filter becomes equivalent to the standard geometric mean. We refer the interested reader to [5] for more details.

Figure 8 shows an illustration of the application of an inhibition-augmented filter that is selective for vertical line endings pointing upwards. Figure 8d shows the output of this filter, and the positions of the strongest local output are marked by crosses in the input image. In this example, this filter only responds strongly at the locations where the positive pattern is present.

Figure 9a shows a data set of line endings with different line widths and orientations. We applied the same configured



**Fig. 8** Illustration of the intermediate computations performed in an inhibition-augmented filter that is selective to vertical line endings pointing upwards. **a** We first convolve the input image (of size  $300 \times 300$  pixels) with a Gabor filter which has a wavelength  $\lambda = 6$  and an orientation  $\theta = 0$ . The three enframed inlay images illustrate (*top two*) the enlarged positive and negative prototype patterns and (*bottom*) the structure of the resulting inhibition-augmented filter. The *red ellipses* represent the preferred wavelengths and orientations of the Gabor filters that provide excitatory input to the concerned filter  $S_f$ , while the *blue ellipses* represent the channels of the Gabor filters that provide inhibitory

input. Then, **b** we blur and shift all thresholded (here  $t_1 = 0.2$ ) Gabor responses appropriately. Next, **c** we compute two weighted geometric means (here  $\sigma' = 11.89$ ), one for the excitatory blurred and shifted Gabor responses and the other one for the inhibitory input. Finally, **d** we calculate the thresholded (here  $t_3 = 0.4$ ) output of the inhibition-augmented filter by subtracting a fraction  $\eta$  (here  $\eta = 1$ ) of the inhibitory response from the excitatory response. The local maxima responses correspond correctly to the three vertical line endings in the input images, which are indicated by the *cross markers* (color figure online)

inhibition-augmented filter to the stimuli in this data set, and the responses of this filter are rendered by a gray level shading of the features (Fig. 9b). The maximum response is reached for the feature that was used as a positive prototype in the configuration process while it also reacts, with less than the maximum response, to line endings that differ slightly in scale and orientation. This example illustrates the selectivity and the generalization ability of the proposed filter.

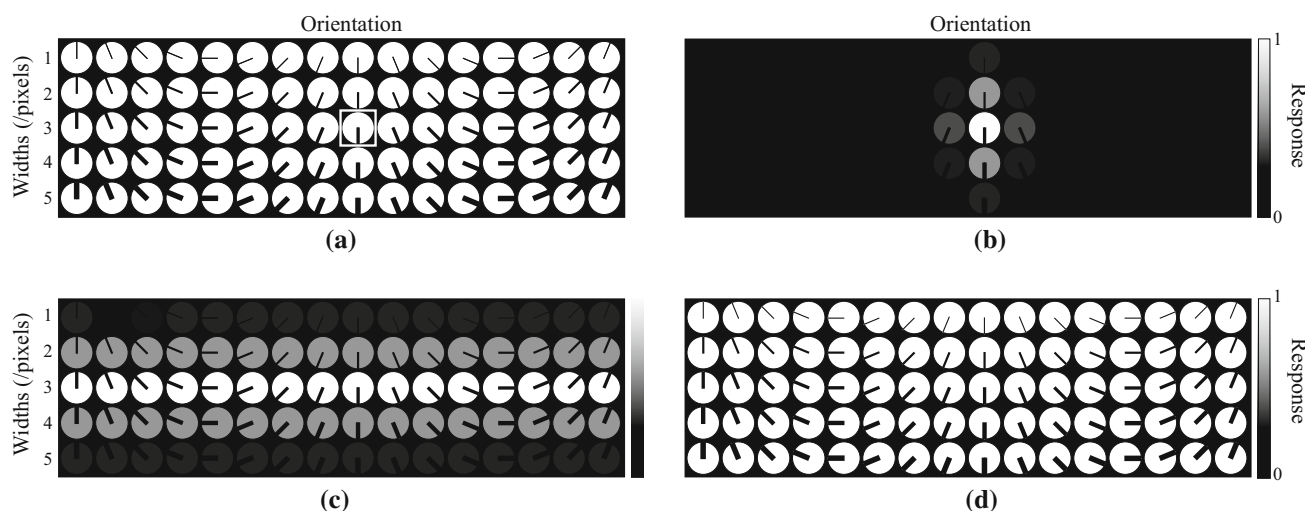
Moreover, in Fig. 10d–f we show the response images of the filter  $S_{\text{big}}$ , which we configured in Sect. 2.4, to the corresponding patterns in Fig. 10a–c. The configured

inhibition-augmented filter correctly responds only to the pattern shown in Fig. 10a but not to the ones in Fig. 10b, c.

## 2.6 Tolerance to geometric transformations

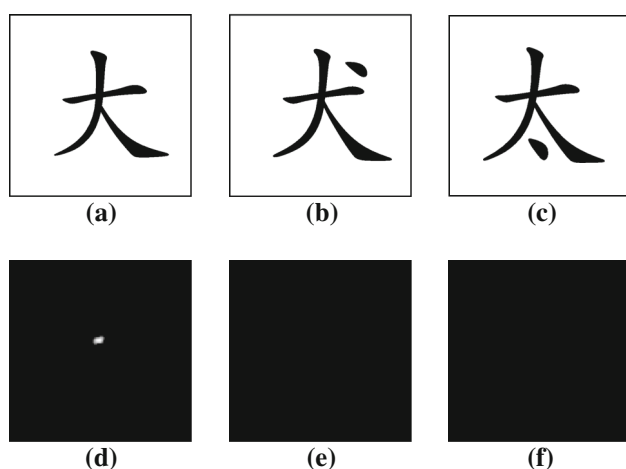
The proposed inhibition-augmented filter can achieve tolerance to scale, rotation and reflection by similar parameter manipulation as proposed for the original COSFIRE filters [5]. Figure 9c, d shows the rotation- and scaling- tolerant responses of the inhibition-augmented filter that correspond to the set of elementary features shown in Fig. 9a. We do not





**Fig. 9** **a** A systematically designed data set of line endings that vary in orientation (in intervals of  $\pi/8$ ) as well as in scale (the line width ranges from 1 pixel to 5 pixels). The enframed feature is the same one shown in Fig. 3b which is used as a positive prototype for configuring an

inhibition-augmented filter. The resulting filter is applied to all features in the data set, and **b** the responses are rendered by shading the features in gray. **c** Rotation-tolerant responses. **d** Rotation- and scale-tolerant results

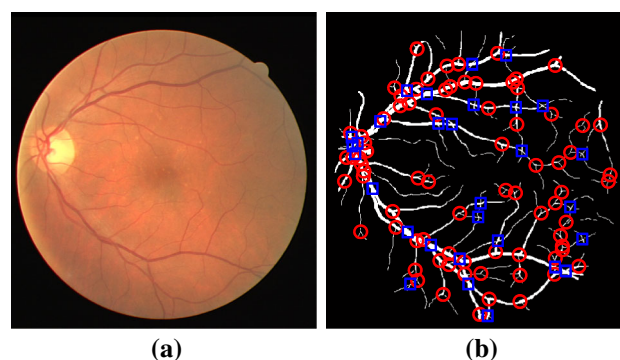


**Fig. 10** **a–c** Images of three similar Chinese letters (of size  $120 \times 120$  pixels). **d–f** The response images of the resulting filter  $S_{big}$  to the corresponding images in (a–c). The filter responds to the letter shown in (a), but does not respond to the letters shown in (b, c)

elaborate on these aspects here, and we refer the reader to [5] for a thorough explanation.

### 3 Applications

In the following, we demonstrate the effectiveness of the proposed inhibition-augmented filter in three practical applications: the detection of vascular bifurcations in retinal fundus images, the recognition of architectural and electrical symbols and the recognition of handwritten digits.



**Fig. 11** Example of a retinal fundus image from the DRIVE data set. **a** Original image (of size  $564 \times 584$  pixels) with filename 21\_training.tif. **b** Binary segmentation of vessels and background (also from DRIVE). The red circles surround vessel bifurcations, blue squares surround crossovers, and this labeling is part of the current work (color figure online)

#### 3.1 Detection of retinal vascular bifurcations

The retina contains cues of the health status of a person. For instance, its vascular geometrical structure can reflect the risk of some cardiovascular diseases such as hypertension [53] and atherosclerosis [14]. The identification of vascular bifurcations is one of the basic steps in such analysis. For a thorough review on retinal fundus image analysis, we refer to [1, 40].

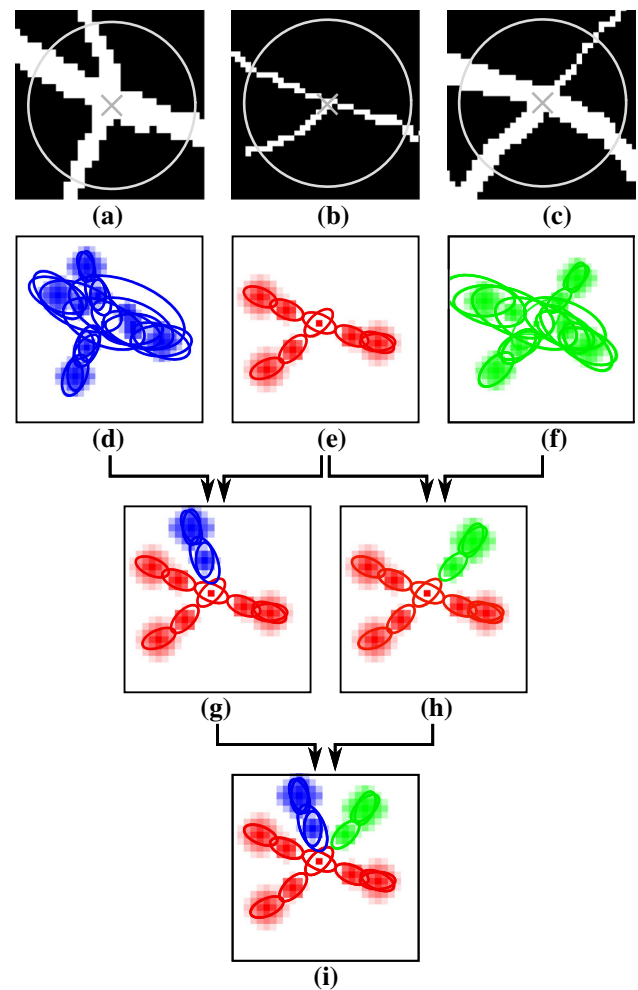
Figure 11 shows an example of a retinal fundus image and its segmentation in blood vessels and background, both of which are taken from the DRIVE data set [48]. It contains 109 blood vessel features (81 bifurcations marked by red circles and 28 crossovers marked by blue squares). A bifurcation-

selective filter configured by the basic COSFIRE approach [5] gives a response also to crossovers and therefore cannot be used to exclusively detect bifurcations. The existing methods that are used to distinguish bifurcations from crossovers preprocess the binary retinal fundus images by morphological operators, such as thinning. Then, they typically apply template matching or connected component labeling, which do not work very well for complicated situations, e.g., two bifurcations that are close to each other can be detected as a crossover. An overview of these methods can be found in [2, 11, 52]. In the following, we illustrate how inhibition-augmented filters that we propose can be configured to detect only vascular bifurcations in retinal fundus images.

First, we select a bifurcation prototype from a given retinal fundus image and use it as a positive example to configure a COSFIRE filter  $P_{f_1}$  that is composed of excitatory vessel segments. For the configuration of this filter, we use three values of the distance  $\rho$  ( $\rho = \{0, 5, 10\}$ ), threshold value  $t_1 = 0.2$ ,  $t_2 = 0.45$ , a bank of symmetric Gabor filters with eight orientations ( $\theta \in \{\frac{\pi i}{8} \mid i = 0 \dots 7\}$ ) and five wavelengths ( $\lambda \in \{4(2^{\frac{i}{2}}) \mid i = 0 \dots 4\}$ ). Figure 12b, e show an enlarged prototype and the corresponding filter structure, respectively. Then, we apply the configured filter  $P_{f_1}$  to all 20 training retinal fundus images (with filenames from 21\_manual1.gif to 40\_manual1.gif) without tolerance to rotation, scale and reflection transformations. We consider the points that characterize crossover patterns and evoke sufficiently strong responses (which is more than a fraction  $\varepsilon$  of the maximum response to the positive pattern, here  $\varepsilon = 0.2$ ). We then use these patterns as negative prototypes. Figure 12a, c show two of the negative prototypes and the structures of the resulting COSFIRE filters are shown in Fig. 12d, f. We generate an inhibition-augmented filter  $S_{f_1}$  by the method proposed in Sect. 2.4. Figure 12d–i shows how two groups of inhibitory line segments are automatically selected by the proposed configuration procedure.

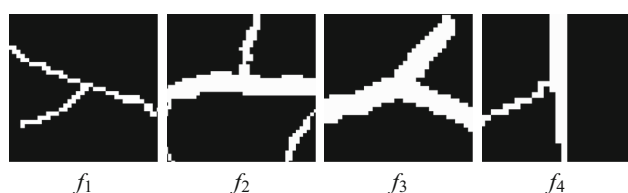
We repeat the above procedure by applying the filter  $P_{f_1}$  in reflection- and rotation-tolerant mode in order to find more negative patterns. Finally, the filter  $S_{f_1}$  contains 19 groups of inhibitory tuples.

The values of the inhibition factor  $\eta$  and the threshold  $t_3$  are determined as follows. We apply the filter  $S_{f_1}$  to the 20 training retinal fundus images and perform a grid search to estimate the best pair of parameters  $\eta$  and  $t_3$ . For  $\eta$ , we consider the range of values  $[0, 5]$ , and for  $t_3$  we consider the range  $[0, 1]$ , both of which are in intervals of 0.01. For each combination of these two parameters, we calculate the precision  $P$  and recall  $R$ . The corresponding harmonic mean ( $2PR/(P + R)$ ) reaches a maximum at an inhibition factor  $\eta = 2$  and threshold  $t_3 = 0.29$  when the precision  $P$  is at least 90%. Here, the filter  $S_{f_1}$  detects 30 bifurcations and achieves 100% precision.



**Fig. 12** Examples of positive and negative prototype patterns. **b** A positive prototype pattern, which is the feature of interest. **a, c** Negative prototype patterns. **d–f** The structure of the filters that are selective for the features in (a–c). **g–h** Two inhibition-augmented filters configured by one positive and one negative prototype. **i** An inhibition-augmented filter configured by one positive and two negative prototypes. The tuples that are indicated by the *red ellipses* come from the positive pattern in (b), and the tuples that are indicated by the *blue* and *green ellipses* come from the negative patterns in (a, c), respectively (color figure online)

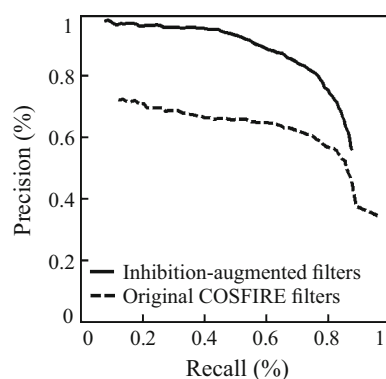
For the remaining bifurcations that are not detected by  $S_{f_1}$ , we perform the following steps. We randomly select one of the undetected bifurcations and use it as a new positive prototype. Then, we use the same procedure as described above to find the inhibitory parts of the new filter  $S_{f_2}$  as well as the corresponding inhibition factor  $\eta$  and threshold value  $t_3$ . The prototype pattern  $f_2$  is shown in Fig. 13. By applying the filters  $S_{f_1}$  and  $S_{f_2}$  ( $\eta(S_{f_2}) = 1.80$ ,  $t_3(S_{f_2}) = 0.37$ ) together, we correctly detect 42 correctly detected bifurcations and no crossovers. We continue increasing the number of filters by using vascular features that are not detected by the previously configured filters. For this given retinal fundus image, we achieve 95% recall and 100% precision with only four



**Fig. 13** A set of four bifurcations ( $f_1 \dots f_4$ ) taken from the DRIVE data set. These four bifurcations are extracted from the binary retinal fundus image shown in Fig. 11b with filename 21\_manual1.gif

**Table 1** Optimal values of  $\eta$  and  $t_3$

	$S_{f_1}$	$S_{f_2}$	$S_{f_3}$	$S_{f_4}$
$\eta$	2.00	1.80	2.00	1.60
$t_3$	0.29	0.22	0.29	0.14



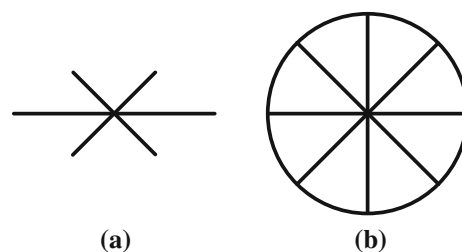
**Fig. 14** Precision–recall plots of the inhibition-augmented method and original COSFIRE method, indicated by the *dashed* and *solid* line, respectively

filters, Fig. 13. Table 1 reports the values of the parameters  $\eta$  and  $t_3$  that were determined with the grid search method described above.

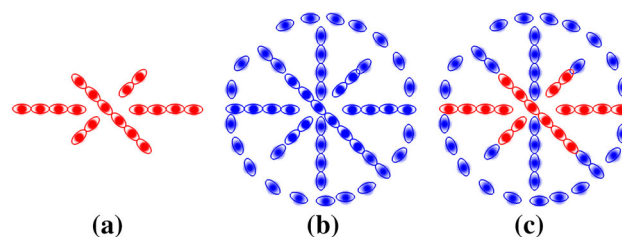
In order to evaluate the performance of proposed approach, we apply the four inhibition-augmented filters to the 20 test retinal fundus images in the DRIVE data set. We perform two experiments with the four filters, the first one using the fine-tuned inhibition factors  $\eta$  and the other one with  $\eta = 0$ . We change the value of the threshold parameter  $t_3(S_{f_i})$  to compute the precision  $P$  and recall  $R$ . For each filter, we alter the threshold value  $t_3(S_{f_i})$  by the same offset value (ranging between  $-0.2$  and  $0.2$  in intervals of  $0.01$ ) which results in the P-R plots shown in Fig. 14. For the same value of recall, the precision of the inhibition-augmented method is substantially higher than that of the method without inhibition.

### 3.2 Recognition of architectural and electrical symbols

Recognition of hand-drawn or scanned architectural and electrical symbols is an important application for the automatic conversion to a digital representation which can then be stored efficiently or processed by auto CAD sys-



**Fig. 15** Example of **a** a symbol that is contained within **b** another symbol



**Fig. 16** **a, b** Structures of COSFIRE filters configured with the positive and negative prototype in Fig. 15a, b, respectively. **c** The structure of the resulting inhibition-augmented filter obtained. The ellipses illustrate the wavelengths and orientations of the selected Gabor filters, and their positions indicate the locations at which their responses are used as input to the concerned COSFIRE filter. The *blobs* within the ellipses represent blurring functions that are used to provide some tolerance regarding the preferred positions. *Red* and *blue ellipses and blobs* indicate Gabor responses that provide, respectively, positive and negative inputs to the inhibition-augmented filter (color figure online)

tems [43,51,55,58]. In the following, we illustrate how the inhibition-augmented filter that we propose is effective for such an application.

We evaluate the proposed approach on the Graphics Recognition Contest (GREC'11) data set [54]. The GREC'11 data set contains 150 different symbol classes, in which the images are of size  $256 \times 256$  pixels. This data set consists of three different sets of images, namely SetA, SetB and SetC. SetA contains 2500 images from 50 symbol classes, SetB comprises 5000 images from 100 classes, and SetC consists of 7500 images from 150 classes. The three data sets contain examples with different scale, rotation and various levels of noise degradation.

In the following, we explain how the proposed inhibition-augmented filters are configured to be exclusively selective for specific symbol classes. Figure 15 shows two such examples of symbol images from the GREC'11 data set. All contour parts of the symbol in Fig. 15a are contained in the symbol in Fig. 15b.

For configuration, we do the following steps. First, we consider a model symbol, such as the one in Fig. 15a, as a positive prototype pattern to configure a COSFIRE filter without inhibition. Figure 16a shows the structure of the resulting filter. Then, we apply the configured filter in rotation- and scale- tolerant mode to all the other 149 model images. We

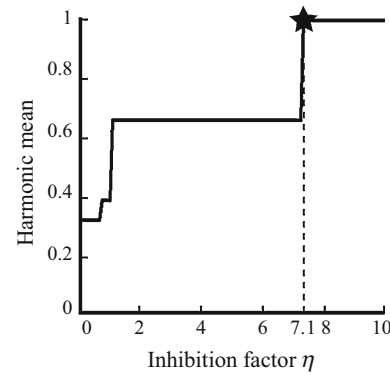
threshold the responses at a given fraction  $\varepsilon$  ( $\varepsilon = 0.3$ ) of the maximal filter response to the positive pattern used for configuration. The other symbol images which evoke strong responses to the filter are considered as negative prototype patterns. For instance, the symbol shown in Fig. 15b is one negative prototype for the pattern in Fig. 15a. The COSFIRE filter structure that corresponds to the pattern in Fig. 15b is shown in Fig. 16b. Next, we compare the structures shown in Fig. 16a, b to identify contour parts to be used for inhibition. In Fig. 16c, we show the structure of the resulting inhibition-augmented filter, in which red and blue ellipses and blobs indicate Gabor responses that provide, respectively, positive and negative inputs to the filter. In this implementation, we consider a bank of Gabor filters with eight orientations ( $\theta \in \{\frac{\pi i}{8} \mid i = 0 \dots 7\}$ ) and two wavelengths ( $\lambda \in \{10, 18\}$ ). We use the empirically determined threshold values  $t_1 = 0.2$  and  $t_2 = 0.5$ . For the blurring function, we use a fixed standard deviation  $\sigma = 4$ . In order to make sure that we extract information from all the line segments of a given prototype, we first use a large set of  $\rho$  values, and then, we remove redundant tuples from the filters as follows. We compute the pairwise dissimilarity proposed in Sect. 2.3 with parameter  $\zeta$  equal to three times the maximum standard deviation of any pair of tuples and delete one tuple from the pair whose dissimilarity value is 0. In this way, the corresponding blurring maps of tuples do not overlap each other.

In order to determine the optimal value of the inhibition factor  $\eta$  for such an inhibition-augmented filter, we perform the following steps. First, we apply the filter to all 150 model symbol images with different values of inhibition factor  $\eta$  in a range between 0 and 10 in interval of 0.1. Then, for each inhibition factor, we calculate the harmonic mean of the precision<sup>2</sup> and recall<sup>3</sup> of this filter. Figure 17 shows the harmonic mean of the concerned filter with different values of inhibition factor. The optimal inhibition factor ( $\eta = 7.1$ ) is the minimum value of  $\eta$  that achieves the highest harmonic mean. In Fig. 17, we indicate this point by a star marker.

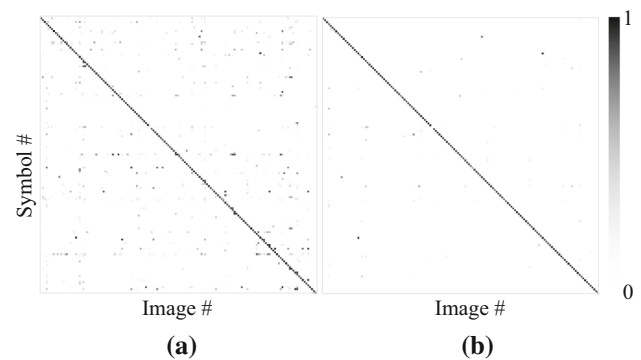
We perform the same procedure on the remaining 149 symbols. We apply the resulting 150 inhibition-augmented filters to the 150 symbol images. Figure 18a, b shows matrices (of size  $150 \times 150$ ) obtained using the COSFIRE filters without inhibition ( $\eta = 0$ ) and the inhibition-augmented COSFIRE filters, respectively. The value of the element  $(i, j)$  in each matrix is the maximum response of the filter configured by symbol  $i$  to symbol image  $j$ . For each filter, we compute the precision and recall. The average precision

<sup>2</sup> We compute precision as the number of images to which the filter correctly responds divided by the total number of images to which the filter responds.

<sup>3</sup> We compute recall as the number of images to which the filter correctly responds divided by the total number of images to which the filter should respond.



**Fig. 17** Harmonic mean of the precision and recall for the filter shown in Fig. 16c with different values of inhibition factor  $\eta$ . The star indicates the minimum inhibition factor ( $\eta = 7.1$ ) that achieves the maximum harmonic mean



**Fig. 18** Matrices (of size  $150 \times 150$ ) obtained by **a** original COSFIRE filters and **b** inhibition-augmented COSFIRE filters with the optimal inhibition factors. The columns represent images, and the rows represent filters which are configured from symbols. The elements of the matrices are the maximum responses of the configured filters to the images, which are rendered as gray levels

achieved by the COSFIRE filters without inhibition is 48.0 % while the one for the inhibition-based filters is 81.7 %. The recall for both methods is 100 %. Compared to the results of the original COSFIRE filters, the matrix obtained by the inhibition-augmented filters is much sparser and the precision is significantly improved.

Before applying the configured filters to the test images in SetA, SetB and SetC, we preprocess each image as follows. We compute the mean value of the intensities of all the pixels in an image. For the images that have a mean intensity value of at least 90 % of the maximum, we apply the morphological operations proposed in [19]. First, we dilate the images by six line-shaped structuring elements of 6-pixel length with different orientations ( $\{0, \frac{\pi}{6}, \frac{\pi}{3}, \dots, \frac{5\pi}{6}\}$ ). Then, we perform a thinning operation followed by six dilations using line-shaped structuring elements of 4-pixel length with equidistant orientations. Finally, we apply opening and thinning followed by a dilation operation using a series of line-shaped structuring elements of 4-pixel length in six ori-



**Table 2** Recognition rates (%) for the three data sets in the GREC'11 set [54]

	SetA	SetB	SetC	Average
Our proposed method	98.80	97.64	98.11	98.18
COSFIRE filters without inhibition [5]	96.80	94.50	91.81	94.37
Spectra of shape context [57]	92.96	90.70	89.79	90.62
Geometric matching [54]	94.76	91.98	85.88	89.39

entations. We do not preprocess the images that have a mean value less than 90 % of the maximum since most of them do not lose part of their contour segments.

We apply the 150 inhibition-augmented filters to each pre-processed image by using the proposed method in rotation- and scaling-tolerant mode with parameters  $\psi = \{\frac{\pi i}{32} \mid i = 0, 1, \dots, 31\}$  and  $v = \{0.5, 0.6, \dots, 2.5\}$ . A given image is classified to the class of the positive prototype symbol by which the inhibition-augmented filter that achieves the maximum response was configured. In Table 2, we compare the results that we achieve with the existing methods on the three data sets. The proposed approach achieves the best results in all data sets.

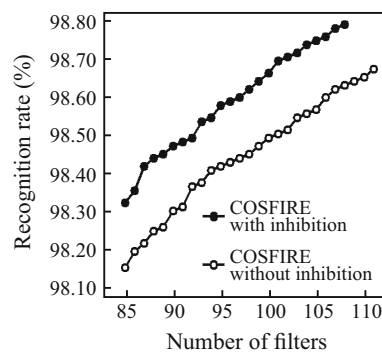
### 3.3 Recognition of handwritten digits

Handwritten digit recognition is an important application in optical character recognition (OCR) systems. Various benchmark data sets and approaches have been proposed, a review of which is given in [33].

In this application, we use the MNIST data set [31] to evaluate the performance of our approach. The data set contains 60,000 training and 10,000 test digit images in gray scale of size  $28 \times 28$  pixels.<sup>4</sup>

For configuration, we randomly select 20 training images from each digit class. We select a random location as the point of interest to configure a COSFIRE filter in each image from the same digit class. The local pattern around such a point should provide at least four tuples to the resulting filter; otherwise, we select another random location. Then, we apply this filter to all the other 180 training images from different digit classes in order to identify negative prototypes. We use the method described in Sect. 2 to configure an inhibition-augmented filter. We repeat the above process for all the 200 training digit images and configure 200 inhibition-augmented filters. In this application, we use a bank of antisymmetric Gabor filters with 16 equidistant orientations ( $\theta \in \{\frac{\pi i}{8} \mid i = 0 \dots 7\}$ ), one wavelength ( $\lambda = 2\sqrt{2}$ )

<sup>4</sup> The MNIST data set is available online: <http://yann.lecun.com/exdb/mnist>.

**Fig. 19** Plots show the recognition rates on the MNIST data set with different numbers of filters by using the methods with and without inhibition

and three values of  $\rho$  ( $\rho \in \{0, 3, 8\}$ ), and threshold their responses with  $t_1 = 0.2$ ,  $t_2 = 0.99$ .

For the application process, we apply these 200 inhibition-augmented filters to 60,000 training images by using the proposed method. We take the maximum response of each filter to these digit images and generate a matrix of size  $60,000 \times 200$ .

Next, we apply a wrapper method for feature selection using support vector machines (SVMs) with a linear kernel. We iteratively add the result of one filter, the one that best improves the sevenfold cross-validation accuracy, and stop the process until no more improvement is achieved. This process results in 108 filters when the 200 inhibition-augmented COSFIRE filters are applied ( $\eta = 1$ ) and 111 filters when the 200 original COSFIRE filters are applied ( $\eta = 0$ ). Then, we use the inhibition-augmented and non-inhibition-augmented training vectors with the selected features to train two multi-class SVMs.

The plots in Fig. 19 show the recognition rates as a function of increasing number of selected filters. The method with inhibition achieves a recognition rate of 98.77 % with 108 filters while the method without inhibition achieves 98.66 % with 111 filters. The inhibition-augmented training vectors of 108 dimensions have 753,019 (11.62 %) zero elements, which is substantially greater than the 277,641 (4.17 %) zero elements in non-inhibition-augmented vectors of 111 dimensions. In this application, the proposed inhibition-augmented COSFIRE filters achieve better recognition rate with less number of filters and with a much sparser representation.

## 4 Discussion

We proposed an inhibition-augmented COSFIRE approach which uses a positive prototype and a set of negative prototypes to configure a filter. The choice of negative prototypes can be either manually specified by a user or automatically



discovered by the system. For instance, the negative prototype shown in Fig. 3, a complete line, is selected by the user. For more complex situations, such as the recognition of symbols and handwritten digits, it is more practical to use an automated process. To discover negative prototypes, we first apply the COSFIRE filter which is configured by a positive prototype pattern to all the other pattern images. The ones which evoke strong responses to the filter are negative prototype patterns.

The response of an inhibition-augmented filter is defined as the difference between the excitatory input and a fraction of the maximum of the inhibitory inputs. The inhibition factor can be adjusted by changing the value of the parameter  $\eta$ . In the detection of vascular bifurcations and the symbol recognition applications, we determine an optimal value of  $\eta$  for each filter as the one that contributes to the maximum harmonic mean on the training images. For the other application, we set the same  $\eta$  value for all filters so that none of them achieves a response to any of the negative patterns.

In neurophysiology, there is an ongoing debate about what kind of neural coding the brain uses to encode the representation of objects. The two extremes in the debate are the grandmother cell theory (i.e., only one specific cell fires for a given pattern) and population coding (i.e., a number of neurons fire for a given pattern with different rates). In the recognition of architectural and electrical symbols, the proposed inhibition-augmented COSFIRE filters work in the way that is similar to the grandmother cell theory. While in the recognition of handwritten digits, they are similar to the population coding. Both applications demonstrate that the inhibition mechanism facilitates sparseness in the representation of information.

The computational cost of the configuration of a COSFIRE filter with inhibition depends on the number of negative prototype patterns and the bank of Gabor filters it uses. An inhibition-augmented filter is configured in less than one second for one positive and one negative prototype pattern with the size of  $512 \times 512$  pixels and a bank of Gabor filters of eight orientations and five wavelengths. The computational cost of the application of an inhibition-augmented filter is proportional to the computations of the excitatory and inhibitory responses and their blurring and shifting operations. For the detection of vascular bifurcations, a retinal fundus image of size  $564 \times 584$  pixels is processed in less than 20 s by four rotation- and reflection-tolerant inhibition-augmented filters. And for the recognition of architectural and electrical symbols, a symbol image of size  $256 \times 256$  pixels is processed in less than 30 s by 150 inhibition-augmented filters without any rotation or scaling tolerances. For the third application, a handwritten digit image of size  $28 \times 28$  pixels is described by 200 inhibition-augmented COSFIRE filters without any rotation or scaling tolerances in less than 5 s.

We used a sequential implementation in MATLAB<sup>5</sup> for all experiments that run on the same standard 3GHz processor.

There are various possible directions for future research. One direction is to apply the proposed inhibition-augmented filters in other objection localization and recognition tasks, as well as image classification. Another direction is to investigate a learning algorithm to determine the output function by assigning different weights to inhibitory and excitatory contour parts.

## 5 Conclusions

The proposed inhibition-augmented filters are versatile trainable keypoint and object detectors as they can be trained with any given positive and negative prototype patterns. We demonstrated the effectiveness of the method with three applications: detection of vascular bifurcations (i.e., without crossovers) in retinal fundus images (DRIVE data set), recognition of architectural and electrical symbols (GREC'11 data set) and the recognition of handwritten digits (MNIST data set). The inclusion of the inhibition mechanism improves the discrimination properties and the performance of COSFIRE filters.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Abramoff, M., Sonka, M.G.M.: Retinal imaging and image analysis. *Biomed. Eng.* **89**, 169–208 (2010)
2. Azzopardi, G., Petkov, N.: Detection of retinal vascular bifurcations by trainable V4-like filters. In: *Computer Analysis of Images and Patterns. Lecture Notes in Computer Science*, vol. 6854, pp. 451–459. Springer, Berlin (2011)
3. Azzopardi, G., Petkov, N.: A CORF computational model of a simple cell that relies on LGN input outperforms the Gabor function model. *Biol. Cybern.* **106**(3), 177–189 (2012)
4. Azzopardi, G., Petkov, N.: A shape descriptor based on trainable COSFIRE filters for the recognition of handwritten digits. In: *Computer Analysis of Images and Patterns (CAIP, York, United Kingdom)*, *Lecture Notes in Computer Science*, vol. 8048, pp. 9–16. Springer, Berlin (2013)
5. Azzopardi, G., Petkov, N.: Trainable COSFIRE filters for keypoint detection and pattern recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(2), 490–503 (2013)

<sup>5</sup> MATLAB scripts for the configuration and application of inhibition-augmented filter can be downloaded from <http://matlabserver.cs.rug.nl>.

6. Azzopardi, G., Petkov, N.: Ventral-stream-like shape representation: from pixel intensity values to trainable object-selective COSFIRE models. *Front. Comput. Neurosci.* **8**(80), 1–9 (2014)
7. Azzopardi, G., Rodríguez-Sánchez, A., Piater, J., Petkov, N.: A push-pull CORF model of a simple cell with antiphase inhibition improves SNR and contour detection. *PLoS ONE* **9**(7), e98424 (2014)
8. Azzopardi, G., Strisciuglio, N., Vento, M., Petkov, N.: Trainable COSFIRE filters for vessel delineation with application to retinal images. *Med. Image Anal.* **19**(1), 46–57 (2015)
9. Bay, H., Ess, A., T, T., Gool, L.V.: Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **110**(3), 346–359 (2008)
10. Beaudet, P.R.: Rotationally invariant image operators. In: *Proceedings of the 4th International Joint Conference on Pattern Recognition*, pp. 579–583 (1978)
11. Bhuiyan, A., Nath, B., Chua, J.J., Ramamohanarao, K.: Automatic detection of vascular bifurcations and crossovers from color retinal fundus images. In: *SITIS*, pp. 711–718. IEEE Computer Society (2007)
12. Bolz, J., Gilbert, C.: Generation of end-inhibition in the visual cortex via interlaminar connections. *Nature* **320**(6060), 362–365 (1986)
13. Brincat, S., Connor, C.: Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat. Neurosci.* **7**(8), 880–886 (2004)
14. Chapman, N., Dell’omo, G., Sartini, M., Witt, N., Hughes, A., Thom, S., Pedrinelli, R.: Peripheral vascular disease is associated with abnormal arteriolar diameter relationships at bifurcations in the human retina. *Clin. Sci.* **103**(2), 111–116 (2002)
15. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005*, vol. 1, pp. 886–893 (2005)
16. Daugman, J.G.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A* **2**(7), 1160–1169 (1985)
17. Fidler, S., Berginc, G., Leonardis, A.: Hierarchical statistical learning of generic parts of object structure. In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, pp. 182–189. IEEE Computer Society (2006)
18. Florack, L., ter Romeny, H.B.M., Koenderink, J.J., Viergever, M.A.: General intensity transformations and differential invariants. *J. Math. Imaging Vis.* **4**(2), 171–187 (1994)
19. Guo, J., Shi, C., Azzopardi, G., Petkov, N.: Recognition of architectural and electrical symbols by COSFIRE filters with inhibition. In: *Computer Analysis of Images and Patterns 2015, Lecture Notes in Computer Science*, vol. 9257, pp. 348–358 (2015)
20. Harris, C., Stephens, M.: A combined corner and edge detection. In: *Proceedings of The Fourth Alvey Vision Conference*, pp. 147–151 (1988)
21. Hubel, D.: *Eye, brain, and vision*, vol. 22. Scientific American, New York (1988)
22. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (London)* **195**, 215–243 (1968)
23. Jones, J., Palmer, L.: An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **58**(6), 1233–1258 (1987)
24. Julesz, B.: Textons, the elements of texture perception, and their interactions. *Nature* **290**(5802), 91–97 (1981)
25. Kalantidis, Y., Tolias, G., Avrithis, Y., Phinikettos, M., Spyrou, E., Mylonas, P., Kollias, S.: Visual image retrieval and localization. *Multimed. Tools Appl.* **51**(2), 555–592 (2011)
26. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: *Pereira, F., Burges, C., Bottou, L., Weinberger, K. (eds.) Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105. Curran Associates, Inc., Red Hook (2012)
27. Kruizinga, P., Petkov, N.: Non-linear operator for oriented texture. *IEEE Trans. Image Process.* **8**(10), 1395–1407 (1999)
28. Lazebnik, S., Schmid, C., Ponce, J.: A sparse texture representation using local affine regions. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1265–1278 (2005)
29. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR ’06*, vol. 2, pp. 2169–2178. IEEE Computer Society (2006)
30. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436–444 (2015)
31. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
32. Li, F.F., Perona, P.: A bayesian hierarchical model for learning natural scene categories. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR ’05*, vol. 2, pp. 524–531. IEEE Computer Society (2005)
33. Liu, C.L., Nakashima, K., Sako, H., Fujisawa, H.: Handwritten digit recognition: benchmarking of state-of-the-art techniques. *Pattern Recognit.* **36**(10), 2271–2285 (2003)
34. Lowe, D.G.: Object recognition from local scale-invariant features. In: *Proceedings of the International Conference on Computer Vision, ICCV ’99*, vol. 2, pp. 1150–1157 (1999)
35. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
36. Mikolajczyk, K., Schmid, C.: Indexing based on scale invariant interest points. In: *Proceedings of the 8th International Conference on Computer Vision*, pp. 525–531 (2001)
37. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(10), 1615–1630 (2005)
38. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
39. Oliva, A., Torralba, A.: Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* **42**, 145–175 (2001)
40. Patton, N., Aslam, T., MacGillivray, T., Deary, I., Dhillon, B., Eikelboom, R., Yogesan, K., Constable, I.: Retinal image analysis: concepts, applications and potential. *Prog. Retin. Eye Res.* **25**(1), 99–127 (2006)
41. Petkov, N., Kruizinga, P.: Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. *Biol. Cybern.* **76**(2), 83–96 (1997)
42. Petkov, N., Visser, W.T.: Modifications of Center-Surround, Spot Detection and Dot-Pattern Selective Operators. *Institute of Mathematics and Computing Science, University of Groningen, The Netherlands (CS 2005-9-01)*, 1–4 (2005)
43. Rebelo, A., Fujinaga, I., Paszkiewicz, F., Maral, A.R.S., Guedes, C., Cardoso, J.S.: Optical music recognition: state-of-the-art and open issues. *IJMIR* **1**(3), 173–190 (2012)
44. Riesenhuber, M., Poggio, T.: Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **2**(11), 1019–1025 (1999)
45. Rolls, E.T., Treves, A.: The relative advantages of sparse versus distributed encoding for associative neuronal networks in the brain. *Netw. Comput. Neural Syst.* **1**(4), 407–421 (1990)
46. Sami, E.B., Mriganka, S.: Response-dependent dynamics of cell-specific inhibition in cortical networks in vivo. *Nat. Commun.* **5**, 5689 (2014)

47. Shi, C., Guo, J., Azzopardi, G., Meijer, J.M., Jonkman, M.F., Petkov, N.: Automatic differentiation of u- and n-serrated patterns in direct immunofluorescence images. In: *Computer Analysis of Images and Patterns 2015*, Lecture Notes in Computer Science, vol. 9256, pp. 513–521. Springer, Berlin (2015)
48. Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imaging* **23**(4), 501–509 (2004)
49. Strisciuglio, N., Azzopardi, G., Vento, M., Petkov, N.: Multiscale blood vessel delineation using B-COSFIRE filters. In: *Computer Analysis of Images and Patterns 2015*, Lecture Notes in Computer Science, vol. 9257, pp. 300–312. Springer, Berlin (2015)
50. Strisciuglio, N., Vento, M., Azzopardi, G., Petkov, N.: Unsupervised delineation of the vessel tree in retinal fundus images. In: *Computational Vision and Medical Image Processing: VIPIMAGE 2015*, vol. 1, pp. 149–155. CRC Press/Balkema, Taylor and Francis Group (2016)
51. Tang, P., Hui, S.C., Fu, C.W.: Online chemical symbol recognition for handwritten chemical expression recognition. In: *2013 IEEE/ACIS 12th International Conference on Computer and Information Science (ICIS)*, pp. 535–540. IEEE (2013)
52. Tsai, C.L., Stewart, C.V., Tanenbaum, H.L., Roysam, B.: Model-based method for improving the accuracy and repeatability of estimating vascular bifurcations and crossovers from retinal fundus images. *IEEE Trans. Inf Technol. Biomed.* **8**(2), 122–130 (2004)
53. Tso, M., Jampol, L.: Path-physiology of hypertensive retinopathy. *Ophthalmology* **89**, 1132–1145 (1982)
54. Valveny, E., Delalandre, M., Raveaux, R., Lamiroy, B.: Report on the symbol recognition and spotting contest. In: *Workshop on Graphics Recognition (GREC 2011)*, vol. 7423, pp. 198–207. Springer (2011)
55. Valveny, E., Dosch, P., Winstanley, A., Zhou, Y., Yang, S., Yan, L., Wenyin, L., Elliman, D., Delalandre, M., Trupin, E., Adam, S., Ogier, J.M.: A general framework for the evaluation of symbol recognition methods. *Int. J. Doc. Anal. Recognit.* **9**(1), 59–74 (2007)
56. Viola, P., Michael, J.J.: Robust real-time face detection. *Int. J. Comput. Vis.* **57**(2), 137–154 (2004)
57. Yang, S.: Spectra of shape contexts: an application to symbol recognition. *Pattern Recognit.* **47**(5), 1891–1903 (2014)
58. Zanibbi, R., Blostein, D., Cordy, J.R.: Recognizing mathematical expressions using tree transformation. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(11), 1455–1467 (2002)
59. Zhang, J., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: a comprehensive study. *Int. J. Comput. Vis.* **73**(2), 213–238 (2007)
60. Zhang, Y., Tian, T., Tian, J., Gong, J., Ming, D.: A novel biologically inspired local feature descriptor. *Biol. Cybern.* **4**(3), 275–290 (2014)

**Ms. Jiapan Guo** received BSc (2009) degree from China Agricultural University, Beijing, China, and MSc (2012) degree in Signal Processing from Xidian University, Xi'an, China. She is currently a PhD candidate at the Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, the Netherlands. Her current research interest is brain-inspired computer vision, including computational visual models of feature and object recognition.

**Mr. Chenyu Shi** received BSc (2009) degree with honor of Technique Specialist Student in Electronic Information Engineering from China Agricultural University (CAU). In 2010, he was selected to join the China-Germany Public Private Partnership project in Bonn University and graduated in Electrification and Automation at CAU in 2011. Currently, he is a PhD candidate at the Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, the Netherlands. His current research interests are in the field of brain-inspired machine vision, which includes computational models of the visual system with applications to contour detection, feature and shape recognition.

**Dr. George Azzopardi** received BSc degree with honors (first class) in Computer Science from Goldsmiths University of London in 2006, and two years later, he received MSc with distinction in Advanced Methods of Computer Science from Queen Mary University of London. In 2013, he received a PhD cum laude from the University of Groningen in visual pattern recognition. Currently, he is Academic Resident (Lecturer) in the Intelligent Computer Systems department of the University of Malta. His current research is in the field of brain-inspired trainable pattern recognition with applications to contour detection, segmentation, feature and shape recognition, as well as predictive modelling of time series data. He was general co-chair of the 16th international CAIP (Computer Analysis of Images and Patterns) conference that was held in Malta in 2015.

**Prof. Nicolai Petkov** received Dr.sc.techn. degree in Computer Engineering (Informationstechnik) from Dresden University of Technology, Dresden, Germany. Since 1991, he is professor of computer science and head of the Intelligent Systems group. He is the author of two monographs and coauthor of another book on parallel computing, holds four patents and has authored and co-authored over 100 scientific papers. His current research is in pattern recognition, machine learning, data analytics and brain-inspired computing, with applications in health care, finance, surveillance, manufacturing, robotics, animal breeding, etc. He is a member of the editorial boards of several journals.